

Algorithmes de recherche exhaustif et guidé pour la recommandation d'un expert dans un réseau professionnel

Maria Malek*,

*LARIS-EISTI

Ave du Parc 95011 Cergy-Pontoise
maria.malek@eisti.fr

1 Introduction

Nous présentons dans cet article un système de recommandation de talents dans un réseau social professionnel. Le réseau est composé d'un ensemble de personnes ayant des liens professionnels. Selon la demande d'un acteur d'origine x , le système doit proposer (recommander) un ou plusieurs acteurs Z répondant au mieux que possible aux critères demandés (exemple : recherche d'une personne ayant des compétences données pour un poste, etc.). Nous proposons un algorithme de recommandation qui utilise les trois types d'information suivants :

- L'information stockée sur la personne (l'acteur ou le nœud du graphe) d'une façon décentralisée au niveau de chaque nœud. Cette connaissance peut être représentée en utilisant une ontologies décrivant les profils utilisateurs.
- L'information décrite par la structure du réseau même. Autrement dit, en explorant les liens partants de l'acteur origine x et en utilisant des algorithmes d'exploration de graphe comme les chemins les moins coûteux, nous pouvons délimiter le champ de recherche de l'ensemble Z (Zhang et Ackerman (2005); Campbell et al. (2003); Adamic et al. (2003); Zhang et al. (2007)). Notre contribution consiste à utiliser l'arbre couvrant du graphe.
- L'information stockée dans les mesures liées aux acteurs intermédiaires passant par les chemins retenus (Newman (2003, 2004); Grossman (2002); Freeman (1979); Everett et Borgatti (1999)). Cette heuristique donne plus d'importance aux chemins ayant des acteurs plus prestigieux.

2 L'algorithme de recommandation

Nous proposons un algorithme ayant l'entrée et la sortie suivantes :

Entrée : une requête posée par l'auteur X formulée par une suite de mots (termes) clés :

Sortie : une suite pondérée d'auteurs $\{(Z_1, P_1), (Z_2, P_2), \dots, (Z_n, P_n)\}$ correspondants au mieux à la requête *ainsi que* : la chaîne sémantique reliant les deux auteurs : une chaîne sémantique reliant deux auteurs X, Z_i est constituée de la liste de mots (termes) clefs se trouvant dans la suite des sommets reliant X à Z_i .

Étapes de l'algorithme

1. Trouver L'arbre couvrant maximum par rapport aux poids des arêtes. Dans notre cas, il s'agit de trouver l'arbre couvrant le plus représentatif.
2. Calculer et stocker les intermédiarités des nœuds.
3. Extraire de l'arbre couvrant une liste de sommets triée à recommander en utilisant l'algorithme exhaustif ou l'algorithme guidé détaillés par la suite.

Le premier algorithme est exhaustif, il est fondé sur la recherche en largeur dans l'arbre couvrant, jusqu'à ce que l'on trouve un auteur à recommander. Le deuxième algorithme utilise l'approche A* pour explorer l'arbre couvrant. Nous définissons une heuristique admissible qui dépend de la similarité entre un profil et la requête ainsi de l'intermédiarité des nœuds-auteurs. Les expériences ont montré que la version guidée trouve souvent la meilleure recommandation tout en améliorant les performances de l'exploration. En comparant les deux algorithmes nous remarquons que L'arbre couvrant n'a pas été recherché en totalité par la version guidée, l'espace de recherche a été réduit de 11% à 49%.

Nous travaillons actuellement sur une élaboration plus fine du profil utilisateur en utilisant une représentation ontologique du domaine. Nous souhaitons également étendre l'algorithme pour des recommandations inter-communautés. Nous réfléchissons également sur l'utilisation de l'arbre couvrant pour des fins sémantiques qui peuvent ramener à la découverte des *rapprochements sémantiques* entre les communautés.

Références

- Adamic, L. A., O. Buyukkokten, et E. Adar (2003). A social network caught in the web. *First Monday* 8(6).
- Campbell, C. S., P. P. Maglio, A. Cozzi, et B. Dom (2003). Expertise identification using email communications. In *CIKM*, pp. 528–531.
- Everett, M. G. et S. P. Borgatti (1999). The centrality of groups and classes. *Journal of Mathematical Sociology* 23(3), 181–201.
- Freeman, L. C. (1979). Centrality in social networks : Conceptual clarification. *Social Networks* 1(3), 215–239.
- Grossman, J. (2002). The evolution of the mathematical research collaboration graph. *Congressus Numeratum*.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review* 45, 167–256.
- Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Science of the United States (PNAS)* 101, 5200–5205.
- Zhang, J. et M. S. Ackerman (2005). Searching for expertise in social networks : a simulation of potential strategies. In *GROUP*, pp. 71–80.
- Zhang, J., M. S. Ackerman, et L. Adamic (2007). Expertise networks in online communities : structure and algorithms. In *WWW '07 : Proceedings of the 16th international conference on World Wide Web*, New York, NY, USA, pp. 221–230. ACM.